# Crystal structure of TNFα complexed with a poxvirus MHC-related TNF binding protein

Zhiru Yang, Anthony P. West, Jr., and Pamela J. Bjorkman

**Supplementary Methods**

**Supplementary References**

**Supplementary Tables**

**Supplementary Figures**

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1683

**Supplementary Methods**

**Protein expression and crystallization.** The gene for 2L was isolated by PCR from the Yaba-like disease virus (ATCC number VR-937). A C-terminally His-tagged version of 2L lacking six C-terminal residues (residues 1-316 of the mature protein) was expressed together with human β2m in baculovirus-infected insect cells (Hi5 cells) using a dicistronic baculovirus transfer vector, pAcUW31 (PharMingen). The 2L and β2m genes used for expression both included native hydrophobic leader sequences. In our numbering system for 2L, the first amino acid of the mature 2L chain is residue 1; a previous numbering system, which started at the initial methionine of the signal peptide, refers to this residue as 17[1-3]. 2L was purified by nickel-NTA and gel filtration chromatography. Although the expression was conducted using methods that resulted in stable heavy chain/β2m heterodimers for other MHC homologs[4], β2m was not detected in purified 2L preparations (data not shown).

A N-terminally His-tagged form of human TNFα (residues 6-157, numbered as in PDB entry 1TNF) was expressed in the *E. coli* cytoplasm using the expression vector pET23a (Novagen), and soluble TNFα was purified from bacterial lysates by nickel-NTA and gel filtration chromatography.

Purified 2L and TNFα trimer were mixed at a 3:1 molar ratio and passed over a S200 16/60 gel filtration column in 20 mM Tris pH 7.5, 150 mM NaCl. Consistent with previous reports[1-3], monomeric 2L lacking β2m bound stably to TNFα. The 2L–TNFα complex was isolated and then concentrated to 14 mg/ml and used to set up


Nature Structural & Molecular Biology: doi:10.1038/nsmb.1683

crystallization trials by hanging drop vapor diffusion. Crystals were obtained in 7-12% PEG 6K, sodium citrate pH 5.0-5.5, 0.2 M NaBr in monoclinic space group $P2_1$ ($a$ = 101.7, $b$ = 170.3, $c$ =122.0, β = 92.04˚) with two 3:1 2L–TNFα complexes per asymmetric unit. The crystal used for data collection had been soaked in 0.5 mM methylmercury chloride, but no heavy atoms were found in the refined structure.

**X-ray crystallography.** Diffraction data were collected at beamline 11-1 of the Stanford Synchrotron Radiation Laboratory; wavelength of data collection was 1.0047 Å. The dataset was collected at 100 K from a single crystal. The data were processed using HKL2000[5] (Supplementary Table 1). The structure was determined by molecular replacement using a TNFα search model (PDB code 1TNF)[6] followed by location of 2L molecules using the coordinates of the α1-α2 platform of ZAG (PDB code 1T7V)[7] in a real space search. After rigid-body refinement[8], rebuilding was done using the program O[9] and $2F_o$-$F_c$ annealed omit maps, alternating with refinement using CNS[10]. Non-crystallographic symmetry (NCS) restraints (300 kcal/mol•Å$^2$) and group B factor refinement were used during refinement. Ordered carbohydrate was attached to 2L residues Asn6, Asn52, and Asn83, but not to Asn197, the remaining potential N-linked glycosylation site in the 2L sequence. The following cysteine pairs were disulfide bonded: 2L 97-163 (difference Fourier maps indicated some radiation-induced breakage of this bond[11]), 202-281, 295-299, and TNFα 69-101. 2L residues 300-316, its C-terminal His-tag, and the N-terminal His-tag of TNFα were disordered. The final model consisted of six NCS-related sets of atoms: 2L residues 1–299, TNFα residues 6–157, 16 water molecules, and 42 atoms of carbohydrate. A Ramachandran plot determined that 80.4%


Nature Structural & Molecular Biology: doi:10.1038/nsmb.1683

of residues are in the most favored region, 14.5% are in the additional allowed region, 4.0% are in the generously allowed region, and 1.0% are in the disallowed region.

For analysis of contacts and buried surface areas, the 2L α1-α2 platform was defined as residues 1-179 and the α3 domain was defined as residues 180-299. The CCP4 program AREAIMOL[12] was used to calculate buried surface areas using a 1.4 Å probe. To identify interacting residues, a maximum distance of 4.0 Å for van der Waals interactions was used. Shape complementarity indices ($S_C$) were calculated as described[13] using the Sc program in CCP4[12]. Groove surface area calculations were performed as previously described[14,15]. Structural superpositions were performed using LSQMAN[16]. Figures were prepared with MOLSCRIPT[17], RASTER-3D[18], and PyMOL[19]. The domain angle displacement for the 2L α3 domain relative to the HLA-A2 α3 domain (~42°) was calculated by measuring the angle between the long axes of each domain after alignment of the α1-α2 platforms of both molecules. The long axes of the α3 domains were defined using the coordinates of 2L residues Gly176 and Asn191 and HLA-A2 residues Glu177 and His192. The second rotation angle (~40°) relating the α3 domains was estimated using the angle defined by 2L Arg270 and Asn191 and HLA-A2 Val248.

**Supplementary References**


1. Brunetti, C.R. et al. *Proc. Natl Acad. Sci. USA* **100**, 4831-6 (2003).
2. Rahman, M.M., Barrett, J.W., Brouckaert, P. & McFadden, G. *J. Biol. Chem.* **281**, 22517-26 (2006).
3. Rahman, M.M. et al. *Virology* **386**, 462-8 (2009).
4. Yang, Z. & Bjorkman, P.J. *Proc. Natl Acad. Sci. USA* **105**, 10095-100 (2008).
5. Otwinowski, Z. & Minor, W. *Meth. Enzymol.* **276**, 307-326 (1997).
6. Eck, M.J. & Sprang, S.R. *J. Biol. Chem.* **264**, 17595-605 (1989).
7. Delker, S.L., West, A.P., Jr., McDermott, L., Kennedy, M.W. & Bjorkman, P.J. *J. Struct. Biol.* **148**, 205-13 (2004).
8. McCoy, A.J. et al. *J. Appl. Cryst.* **40**, 658-674 (2007).
9. Jones, T.A. & Kjeldgaard, M. *Methods Enzymol.* **277**, 173-208 (1997).
10. Brünger, A.T. et al. *Acta Cryst. D* **54**, 905-921 (1998).
11. Burmeister, W.P. *Acta Crystallogr. D* **56**, 328-41 (2000).
12. CCP4. *Acta Crystallogr.* **D50**, 760-763 (1994).
13. Lawrence, M.C. & Colman, P.M. *J. Mol. Biol.* **234**, 946-950 (1993).
14. Lebron, J.A. et al. *Cell* **93**, 111-23 (1998).
15. Zeng, Z. et al. *Science* **277**, 339-45 (1997).
16. Kleywegt, G.J. *Acta Crystallogr. D* **52**, 842-57 (1996).
17. Kraulis, P.J. *J. Appl. Crystallogr.* **24**, 946-950 (1991).
18. Merritt, E.A. & Murphy, M.E. *Acta Crystallogr. D* **50**, 869-73 (1994).
19. DeLano, W.L. *The PyMOL Molecular Graphics System.*, (DeLano Scientific, San Carlos, California,, 2002).



Nature Structural & Molecular Biology: doi:10.1038/nsmb.1683

20. Madden, D.R. *Annu. Rev. Immunol.* **13**, 587-622 (1995).

21. Li, P. et al. *Nat. Immunol.* **2**, 443-51 (2001).

22. Banner, D.W. et al. *Cell* **73**, 431-45 (1993).

**Supplementary Table 1.** Data collection and refinement statistics

| | **2L–TNFα** |
|---|---|
| **Data collection** | |
| Space group | $P2_1$ |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 101.7, 170.3, 122.0 |
| *β* (°) | 92.04 |
| Resolution (Å) | 50.0-2.80 (2.90-2.80)* |
| $R_{sym}$ or $R_{merge}$ | 8.1 (47.2) |
| $I / \sigma I$ | 16.5 (2.3) |
| Completeness (%) | 96.8 (81.0) |
| Redundancy | 3.7 (3.4) |
| | |
| **Refinement** | |
| Resolution (Å) | 50.0-2.80 |
| No. reflections | 98697 |
| $R_{work}$ / $R_{free}$ | 23.6 / 26.6 |
| No. atoms | |
| Protein | 21660 |
| Ligand/ion | 252 |
| Water | 96 |
| *B*-factors | |
| Protein | 74.0 |
| Ligand/ion | 108.7 |
| Water | 48.4 |
| R.m.s. deviations | |
| Bond lengths (Å) | 0.007 |
| Bond angles (°) | 1.48 |

Data were collected on a single crystal, and in the refinement, 5% of unique reflections were removed as a test set for $R_{free}$ calculation. *Values in parentheses are for highest-resolution shell.

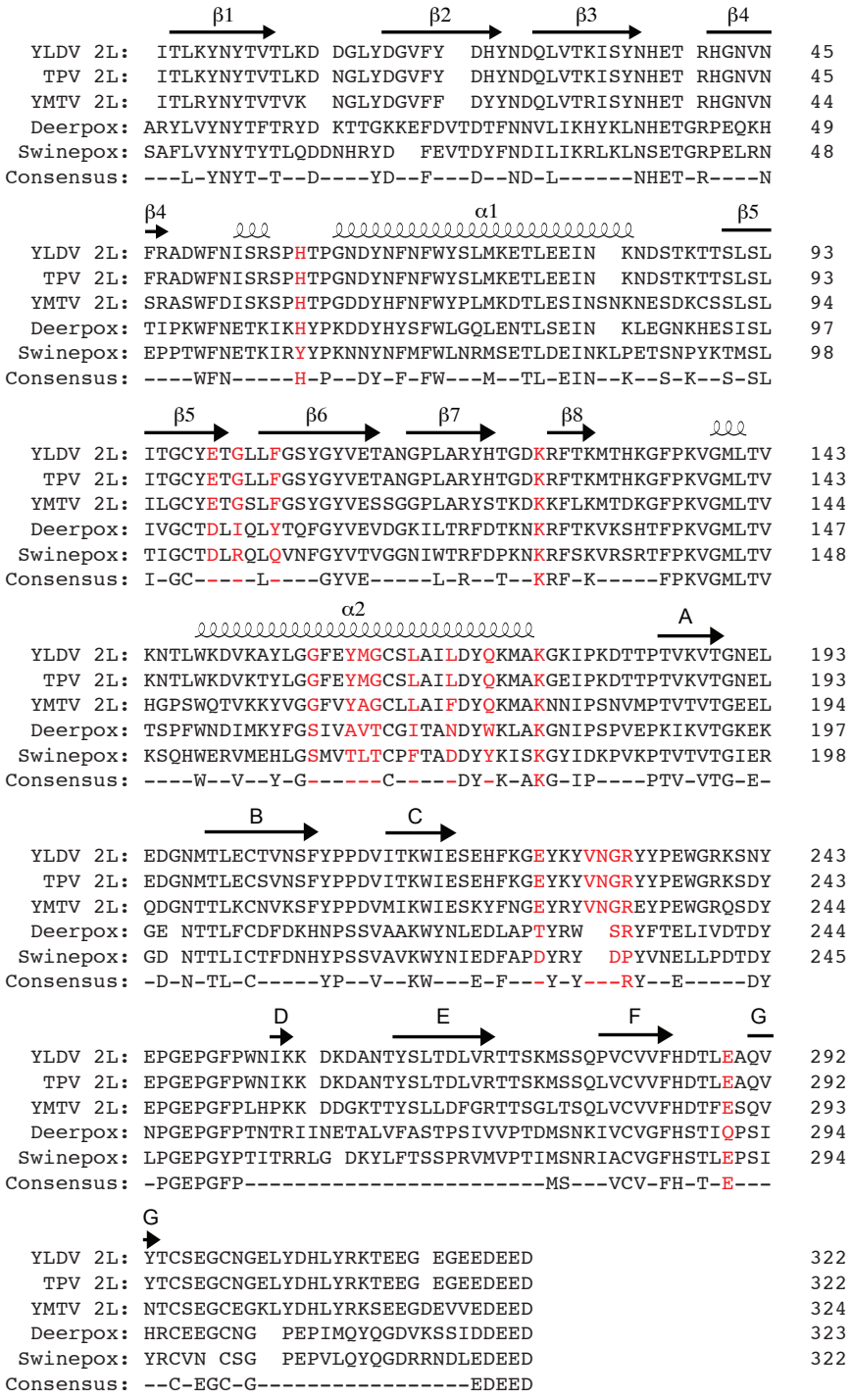
**Supplementary Table 2.** Root mean square deviations (RMSD) for superpositions of the α1-α2 platform of 2L with platforms of MHC and MHC-like molecules.

| Protein | RMSD | Cαs aligned | PDB code | ref. |
|---|---|---|---|---|
| HLA-A2 | 1.43 Å | 104 atoms | 2CLR | 1 |
| M10.5 | 1.45 Å | 105 atoms | 1ZS8 | 2 |
| UL18 | 1.56 Å | 105 atoms | 3D2U | 3 |
| ZAG | 1.69 Å | 107 atoms | 1ZAG | 4 |
| T22 | 1.70 Å | 96 atoms | 1C16 | 5 |
| FcRn | 1.73 Å | 123 atoms | 1EXU | 6 |
| CD1d1 | 1.75 Å | 113 atoms | 1CD1 | 7 |
| m153 | 1.75 Å | 75 atoms | 2O5N | 8 |
| HFE | 1.77 Å | 102 atoms | 1A6Z | 9 |
| m144 | 1.81 Å | 90 atoms | 1U58 | 10 |
| m157 | 1.83 Å | 80 atoms | 2NYK | 11 |
| MIC-A | 1.85 Å | 103 atoms | 1HYR | 12 |
| RAE-1 | 1.91 Å | 91 atoms | 1JFM | 13 |


1. Collins, E.J., Garboczi, D.N. & Wiley, D.C. *Nature* **371**, 626-9 (1994).
2. Olson, R., Huey-Tubman, K.E., Dulac, C. & Bjorkman, P.J. *PLoS Biol.* **3**, e257 (2005).
3. Yang, Z. & Bjorkman, P.J. *Proc. Natl Acad. Sci. USA* **105**, 10095-100 (2008).
4. Sanchez, L.M., Chirino, A.J. & Bjorkman, P. *Science* **283**, 1914-9 (1999).
5. Wingren, C., Crowley, M.P., Degano, M., Chien, Y. & Wilson, I.A. *Science* **287**, 310-4 (2000).
6. West, A.P., Jr. & Bjorkman, P.J. *Biochemistry* **39**, 9698-708 (2000).
7. Zeng, Z. et al. *Science* **277**, 339-45 (1997).
8. Mans, J. et al. *J Biol Chem* **282**, 35247-58 (2007).
9. Lebron, J.A. et al. *Cell* **93**, 111-23 (1998).
10. Natarajan, K. et al. *J Mol Biol* **358**, 157-71 (2006).
11. Adams, E.J. et al. *Proc Natl Acad Sci U S A* **104**, 10128-33 (2007).
12. Li, P. et al. *Nat. Immunol.* **2**, 443-51 (2001).
13. Li, P., McDermott, G. & Strong, R.K. *Immunity* **16**, 77-86 (2002).

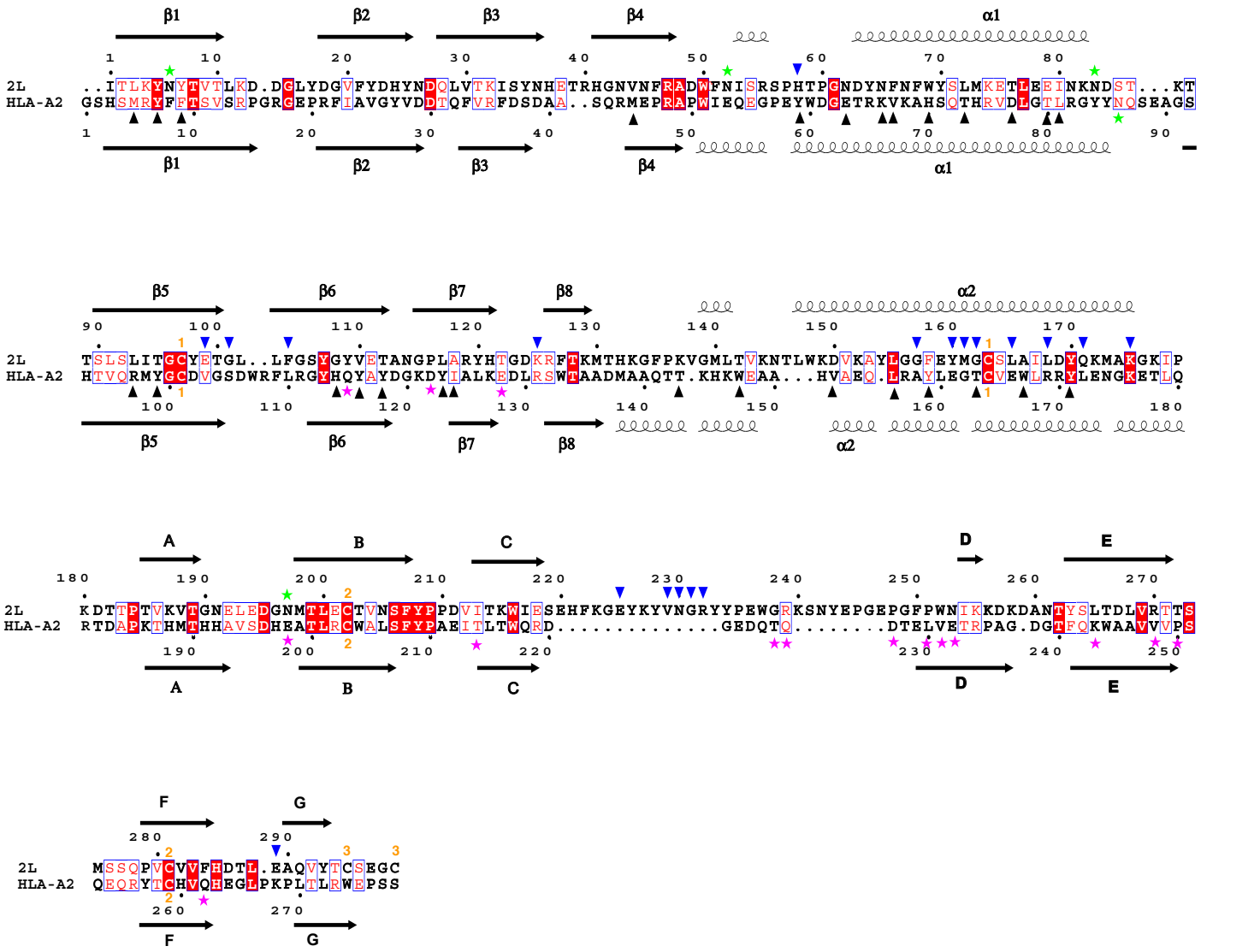
**Supplementary Table 3.** Interactions between 2L and TNFα. % ASA (accessible surface area) = percent of the total interface ASA contributed by each residue. Residues contributing less that 2% of interface ASA were excluded unless they are involved in hydrogen bonds or salt bridges (indicated by asterisks).

| 2L | | | TNFα | | |
|---|---|---|---|---|---|
| Residue | Location | %ASA | Residue | (subunit) | Location |
| HIS58* | α1 helix | 4 | TYR87 | (B) | D-E loop |
| GLU99* | β5 strand | 6 | ARG31 | (A) | A-A′ loop |
| GLU99* | β5 strand | | ARG32 | (A) | A-A′ loop |
| GLY101 | β5-β6 loop | 2 | ARG31 | (A) | A-A′ loop |
| PHE104 | β6 strand | 6 | ASN30 | (A) | A-A′ loop |
| PHE104 | β6 strand | | ARG31 | (A) | A-A′ loop |
| LYS125 | β7-β8 loop | 5 | SER86 | (B) | D-E loop |
| GLY157 | α2 helix | 3 | SER86 | (B) | D-E loop |
| GLY157 | α2 helix | | TYR87 | (B) | D-E loop |
| TYR160 | α2 helix | 9 | ARG82 | (B) | D strand |
| TYR160 | α2 helix | | SER86 | (B) | D-E loop |
| TYR160 | α2 helix | | GLU127 | (B) | F-G loop |
| MET161 | α2 helix | 8 | ASN34 | (A) | A-A′ loop |
| MET161 | α2 helix | | ARG82 | (B) | D strand |
| MET161 | α2 helix | | VAL91 | (B) | E strand |
| GLY162 | α2 helix | 2 | TYR87 | (B) | D-E loop |
| LEU165 | α2 helix | 3 | ALA33 | (A) | A-A′ loop |
| LEU165 | α2 helix | | VAL91 | (B) | E strand |
| LEU168 | α2 helix | 3 | ARG32 | (A) | A-A′ loop |
| GLN171 | α2 helix | 4 | GLN21 | (A) | A-A′ loop |
| LYS175 | α2 helix | 4 | GLN21 | (A) | A-A′ loop |
| GLU225* | α3 insertion | 2 | ARG44 | (A) | B′ strand |
| VAL229 | α3 insertion | 5 | GLU42 | (A) | B′ strand |
| VAL229 | α3 insertion | | ARG44 | (A) | B′ strand |
| ASN230 | α3 insertion | 7 | ASN30 | (A) | A-A′ loop |
| ASN230 | α3 insertion | | LEU37 | (A) | A′ strand |
| ASN230 | α3 insertion | | LEU43 | (A) | B′ strand |
| GLY231 | α3 insertion | 6 | LEU37 | (A) | A′ strand |
| GLY231 | α3 insertion | | VAL41 | (A) | B′ strand |
| GLY231 | α3 insertion | | GLU42 | (A) | B′ strand |
| ARG232* | α3 insertion | 5 | GLU42 | (A) | B′ strand |
| ARG232 | α3 insertion | | ARG44 | (A) | B′ strand |
| GLU289 | F-G loop | 4 | GLN27 | (A) | A-A′ strand |
| GLU289* | F-G loop | | ARG31 | (A) | A-A′ strand |

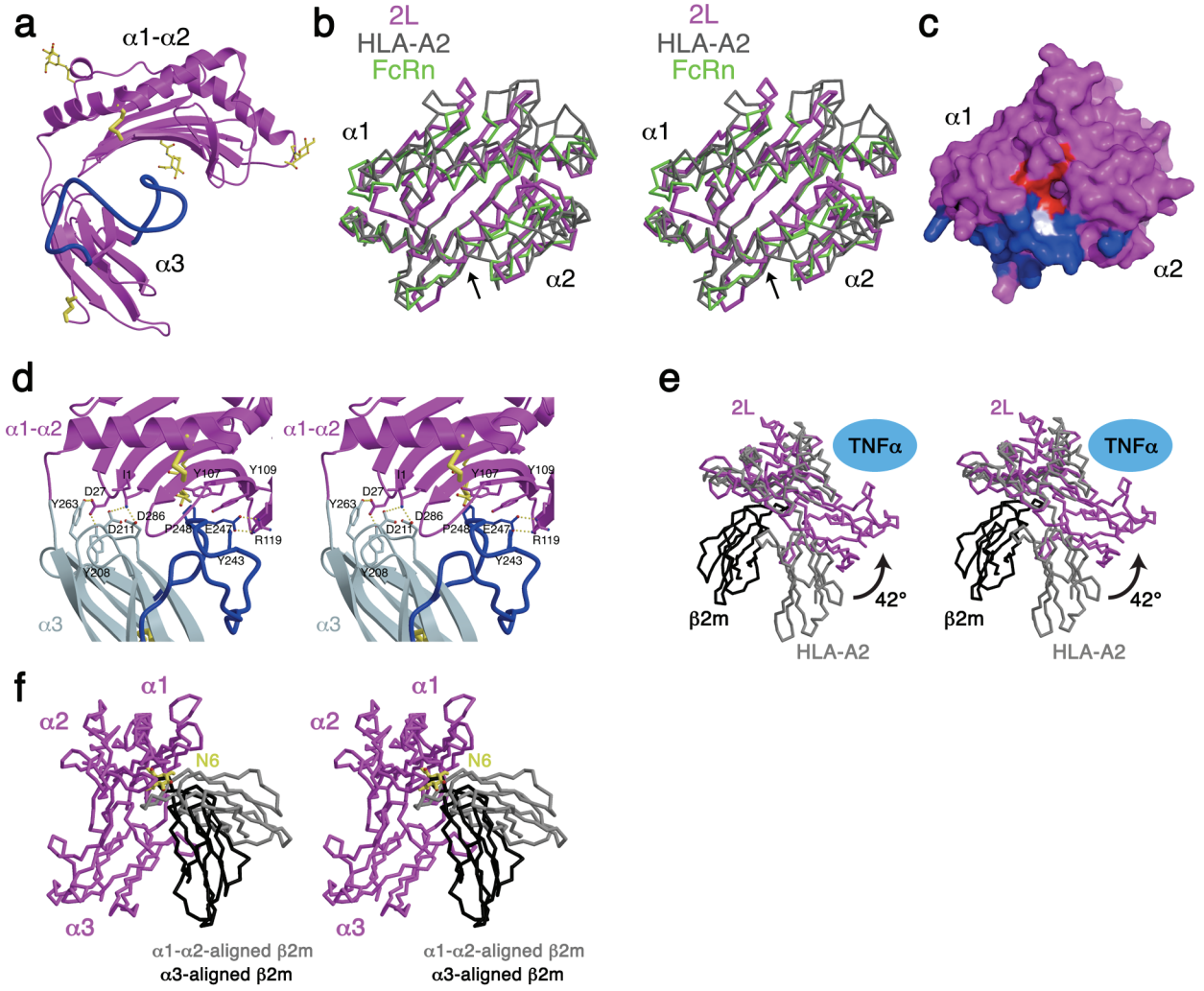



**Supplementary Figure 1. Sequence alignment of poxvirus 2L proteins.** YLDV = Yaba-like disease virus; TPV = Tanapox virus; YMTV = Yaba monkey tumor virus; Deerpox = Deerpox virus strain W-848-83; Swinepox = Swinepox open reading frame SPV003. Consensus amino acid is shown when identity is >70%. Secondary structures are indicated as springs for α-helices and arrows for β-strands. Residues colored red indicate that the amino acid at this position contacts TNFα in the YLDV 2L–TNFα complex structure.The different 2L proteins show varying degrees of species specificity. For example, swinepox 2L binds tightly to porcine TNFα, but not to human TNFα, and porcine TNFα does not bind well to TPV 2L (which has a high affinity for human TNFα). Compensating differences in residues at the 2L–TNFα interface, such as the substitution of TPV 2L residues Gly231-Arg232 for asparate and proline in swinepox 2L, and

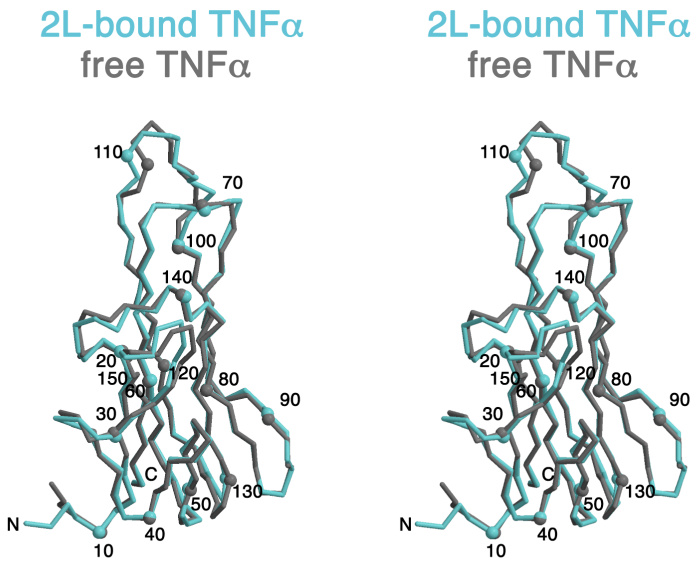
human TNFα residue Glu42 for a lysine in porcine TNFα (Supplementary Fig. 5), likely account for species-specific binding preferences. Consistent with the varying degrees of specificity of poxvirus 2L proteins for binding TNFα from different species, the TNFα binding site on 2L (indicated in red) is less conserved than the rest of the protein: 26% (5 of 19) of the binding site residues are conserved in 2L proteins from four or more poxviruses, compared with an overall conservation of 42% (125 of 299 residues). Sequence diversity at the binding site is likely related to adaptation of poxviruses to the TNFα of their host species. Variability in classical class I MHC molecules primarily maps to a different location, the peptide-binding groove[20].



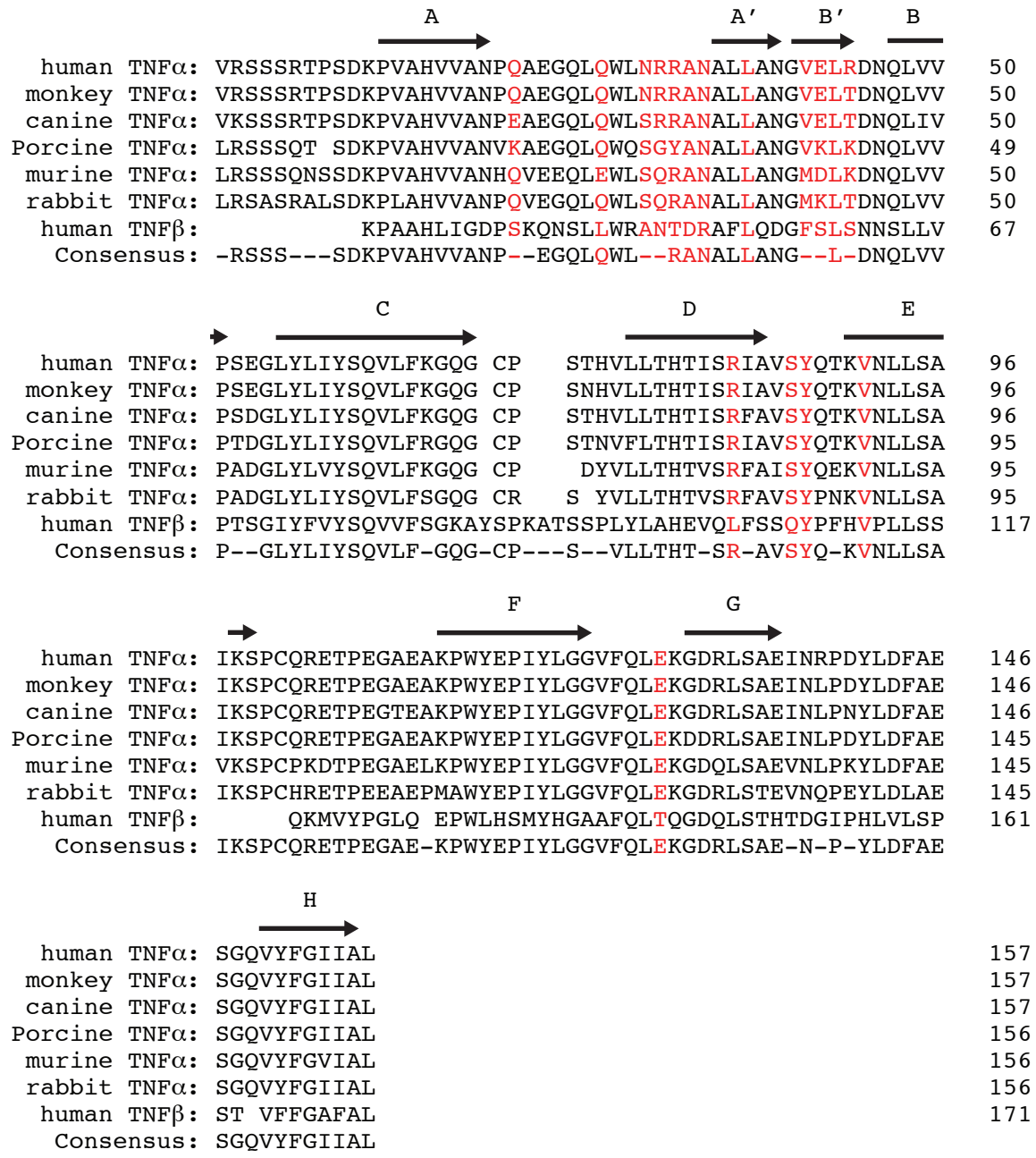

**Supplementary Figure 2. Structure-based sequence alignment of 2L and the class I MHC molecule HLA-A2.** Secondary structures are indicated as springs for α-helices and arrows for β-strands. Conserved residues are shown as white letters in red boxes, and conservative substitutions are indicated as red letters. 2L residues that contact TNFα are indicated with downward-pointing blue triangles. Residues in the HLA-A2 peptide-binding groove are indicated with upward-pointing black triangles. Glycosylation sites are indicated with green stars. Disulfide bonds are indicated with the same orange numeral above or below bonded cysteines. Residues critical for class I MHC binding to the T cell co-receptor CD8 are indicated below the HLA-A2 sequence as purple stars. A comparison with 2L indicates that it does not contain the canonical CD8 binding site found on class I MHC proteins.




**Supplementary Figure 3. Comparison of 2L and MHC structures. a**, Ribbon diagram of 2L with the α3 domain insertion highlighted in dark blue. **b**, Stereo Cα superposition of the α1-α2 platforms of 2L (magenta) with HLA-A2 (gray) and rat FcRn (green). Groove closure in 2L is partially accomplished by displacement of the N-terminal half of each helix towards the opposite helix. Although the 2L α2 domain helix includes a similar kink to that observed in the α2 domain helices of FcRn and CD1 (indicated on the 2L structure with an arrow), 2L lacks the proline found near this kink in other MHC homologs (residues at this position: 2L Ser164/FcRn Pro166/CD1a Pro167/ HLA-A2 Val165). **c**, A surface representation of the 2L α1-α2 platform shown in the same orientation as in panel b. Residues contributing to the 60 $Å^2$ of accessible groove surface area are red, the TNFα binding site is blue, and one residue belonging to both sets is light blue. The surface includes a small pocket (to the immediate right of the red residues) analogous to a pocket in the MIC-A structure[21]. **d**, Stereo close-up of the interface between the 2L α1-α2 platform (magenta) and α3 domain (light blue, with dark blue highlighting the insertion in the α3 domain). Sidechains of residues involved in stabilizing the interface are highlighted and labeled. **e**, Stereo superposition of 2L (magenta) and HLA-A2 (gray, with the β2m light chain in black). An arrow indicates the displacement of the 2L α3 domain compared with the HLA-A2 α3 domain. The approximate position of TNFα is indicated as an oval. **f**, Stereo representation of 2L (magenta) with β2m docked in two possible positions (gray and black). The gray β2m is docked in the position it would occupy if it interacted with the bottom of the 2L α1-α2 domain as β2m interacts with the HLA-A2 α1-α2 region. The black β2m is docked in the position it would occupy if it interacted with the 2L α3 domain as β2m interacts with HLA-A2 α3. The N-linked carbohydrate attached to 2L Asn6 is shown as yellow sticks.




**Supplementary Figure 4. Stereo superposition of 2L-bound TNFα with free TNFα.** A single subunit of the TNFα trimer is shown. The 2L-bound TNFα trimer resembles its counterparts in structures of unliganded TNFα[6] and TNFR1-bound TNFβ[22]: r.m.s. deviations of 0.91 Å (148 Cαs) and 0.95 Å (134 Cαs) for the TNFα-TNFα and TNFβ-TNFα comparisons. The relatively modest conformational changes in 2L-bound TNFα include movement of the A-A' loop (residues 31-35) at the 2L binding site, which permits two salt bridge interactions: 2L Glu99 with TNFα Arg31/Arg32, and 2L Glu289 with TNFα Arg31. The corresponding region of TNFβ, residues 48-52, also exhibits changes in mainchain and sidechain torsion angles upon binding TNFR1[22]. Smaller changes were also observed in the other binding site loops of TNFα, including residues 144-148 and 21-25.




**Supplementary Figure 5. Sequence alignment of mammalian TNFα proteins and human TNFβ.** Secondary structures are indicated as arrows for β-strands. Residues colored red indicate that the amino acid at this position contacts 2L in the 2L–human TNFα complex structure. Consensus amino acid is shown when identity is >70%.